

algorithm. Associations with health, socio-demographic and work characteristics were explored using Spearman correlations. Linear regression was performed to identify correlates. Ceiling effects were analysed. For comparative purposes, normative SF-36v1 data from a general employed population were obtained from Release 11 of the Household, Income and Labour Dynamics in Australia (HILDA) survey (n=11234). **RESULTS:** Mean (SE) health utility was higher in males 0.794 (0.004, n=927) than females 0.773 (0.003, n=2387). Health utility was negatively correlated with psychological distress (K10) (r= -0.69 females, r= -0.69 males), comorbidity count (r= -0.35 females, r= -0.28 males), and effort-reward imbalance (ERI) (r= -0.36 females, r= -0.33 males). SF-6D was independent of age, BMI and annual salary. Lower SF-6D was significantly associated (p<0.05) with higher comorbidities, K10, ERI, education (in females), absenteeism, and lower physical activity (in females). The Australian normative mean (SE) SF-6D was 0.792 (0.002, n=4955) for males and 0.775 (0.003, n=4639) for females. Correlates and associations were similar in the normative sample, which showed an additional significant inverse association with age and positive association with salary in the general employed. **CONCLUSIONS:** SF-6D health utilities differentiated between and associated as expected with key health, social and economic factors. These results validate the suitability of SF-6D to measure perceived health states in an employed population.

PRM28

A COMPARISON OF THREE LANGUAGE VERSIONS OF THE EQ-5D-5L DIMENSION SCALES IN SINGAPORE

Luo N¹, Wang Y¹, How CH², Wong KY¹, Tay EG², Thumboo J³

¹Saw Swee Hock School of Public Health, National University of Singapore, Singapore, ²SingHealth Polyclinics, Singapore, ³Singapore General Hospital, Outram, Singapore

OBJECTIVES: To investigate the comparability of the English, Chinese and Malay versions of the 5 EQ-5D-5L dimension scales in Singapore. **METHODS:** A cross-sectional survey was conducted among visitors to a public primary care institution in Singapore. Participants were interviewed face-to-face in a language they preferred to measure the severity of the EQ-5D-5L response labels ('no', 'slight', 'moderate', 'severe', and 'unable to (do)/extreme' for each dimension scale) they perceived, using a 0-100 numeric rating scale. Participants were also asked to describe 25 dimension-specific health scenarios (5 for each dimension), each using one of the EQ-5D-5L response labels. Differences in label ratings and selection between participants speaking different languages were assessed using regression analysis. The individual-level rank order of label ratings was examined to assess violation of scale ordinality. **RESULTS:** Severity of the EQ-5D-5L response labels perceived by Chinese-speaking (n=256) and English-speaking participants (n=257) was similar. Malay-speaking participants (n=230) perceived the 'slight' labels as more severe (adjusted differences in severity score: 10.5 to 14.5) but the 'unable to (do)/extreme' labels as less severe than English-speaking participants (adjusted differences in severity score: -13.3 to -11.0) (p < 0.001 for all). While no major violation of ordinality was observed among English- and Chinese-speaking participants, a noticeable proportion of Malay-speaking participants rated 'slight' as more severe than 'moderate' (up to 29.6%) and 'severe' as more severe than 'unable to (do)/extreme' (up to 33.9%). In label selection exercise, Chinese-speaking participants tended to use less severe labels than English-speaking participants; Malay-speaking participants tended to use 'no' and 'unable to (do)/extreme'. **CONCLUSIONS:** Differences in perception and usage of response labels exist among Singaporeans using different language versions of the EQ-5D-5L dimension scales. Future studies are needed to investigate whether the differences affect the cross-cultural measurement equivalence of the instrument and, if so, how to alleviate the differences.

PRM29

VALIDATION OF RUSSIAN VERSION OF HEALTH UTILITY INDEX QUESTIONNAIRE IN CHILDREN WITH CYSTIC FIBROSIS

Simonova O, Gorinova Y, Vinyarskaya I, Chernikov V

FSBI Scientific Center of Children's Health, RAMS, Moscow, Russia

OBJECTIVES: Quality of life is an important parameter in the evaluation of quality and outcome of health care and treatment, especially in patients with chronic disorders. The aim of this study was to assess the validity and reliability of Russian version of health utility index questionnaire (HUI) in children with cystic fibrosis (CF). **METHODS:** 65 CF children 6-17 years were interviewed using HUI version 15Q, children over 12 years responded on their own (self assessment), health status of children younger 12 years was obtained from parents (proxy assessment). In addition all parents were interviewed using PedsQL questionnaire. Version translated into Russian language and has been adapted. QOL assessment was carried out on systems HUI 2 and 3. Single-attribute utility scores and multi-attribute utility scores evaluated corresponding to each system, ranges from 0 to 1. Forced expiratory volume in 1 s in percentage of predicted (FEV1%) was included as measure of health status. Patients were interviewed twice with a difference of 3 months. **RESULTS:** Reliability was calculated using Chronbach's α which ranges from 0,71 to 0,80 in HUI 2 system and 0,71 to 0,77 in HUI 3 system. Significant difference between patients with FEV1 > 80% and < 80% were revealed in such utilities like: Mobility, Ambulation, Emotions, Pain (p<0,05). Correlations of HUI system attributes including Sensation (r=0,48), Self-care (r=0,32), Cognition (r=0,24) and Dexterity (r=0,38) with PedsQL scale were obtained. Mean value of several utilities significantly increased in 3 month: Sensation, Emotions, Cognition and Pain (p<0,05). **CONCLUSIONS:** Russian version of HUI questionnaire is a reliable and valid instrument for measuring the health-related quality of life in children with CF.

PRM31

A SURVEY OF KNOWLEDGE AND ATTITUDE OF MENOPAUSE AMONG POST-MENOPAUSAL WOMEN IN PAKISTAN

Malhi TH¹, Qadir MI², Khan YH¹, Khan AH³, Adnan AS⁴

¹Universiti Sains Malaysia, Penang, Malaysia, ²GC University, Faisalabad, Pakistan, Faisalabad, Pakistan, ³Universiti Sains Malaysia, Minden, Malaysia, ⁴CKD Resource Centre, Hospital Universiti Sains Malaysia, Kota Bharu, Kelantan, Malaysia

OBJECTIVES: Menopause is a nettlesome phase of women life. A prior knowledge of menopausal changes will not only help women to cope up with this change but it will also increase their quality of life. Current study aims to analyze knowledge and attitudes of postmenopausal women towards menopause. **METHODS:** A cross-sectional survey was conducted from November, 2011 to September, 2012 among representative sample of women aged 45-60 years belonging to three major cities of Punjab, Pakistan. A team of five trained researchers distributed questionnaire to participants via convenience sampling after obtaining consent. Total of 1000 participants were approached (200 via each researcher). 783 completed the questionnaire making response rate of study as 78%. **RESULTS:** 78% and 22% respondents have natural menopause and surgically induced menopause respectively. The mean knowledge and attitude score of participants was 45.13 \pm 9.23 and 52.32 \pm 11.21 respectively. 48% of the participants had poor, 35% had moderate and 17 % had good knowledge regarding awareness, sign and symptoms and causes of menopause. Majority (69.12 %) of participants had positive attitude towards menopause except use of HRT therapy. 51.7 % patients had good socioeconomic status and had good knowledge compared to 49.3 % of participants with poor socioeconomic status (p=0.031). A higher education level (p=0.041), good socio-economic status (p=0.012), and surgically induced menopause (0.032) were the demographic factors that showed statistically significant association with higher knowledge and positive attitude towards menopause. **CONCLUSIONS:** The knowledge of menopausal women regarding signs, symptoms, causes and treatments of menopause was poor in two cities (Faisalabad, Lahore) as compared to Islamabad. It might be due to good socioeconomic status and higher education level of participants residing in Islamabad. Health care professionals and authorities should take initiatives to aware menopausal women regarding menopause so that they can cope up better with this phase of life.

RESEARCH ON METHODS – Statistical Methods

PRM32

ASSESSING STATISTICAL METHODS FOR CAUSAL INFERENCE IN OBSERVATIONAL DATA

Parks DC, Lin X, Lee KR
GlaxoSmithKline, Collegeville, PA, USA

OBJECTIVES: In observational studies, subjects are assigned to treatment groups without the benefits of randomization, resulting in potential bias in the estimation of the treatment effect. We assess the performance of 5 different statistical methods used for bias correction and causal inference under different conditions – multivariate regression (MR), propensity score matching (PSM), propensity score stratification (PSST), doubly robust estimation (DR) and inverse probability treatment weighting (IPTW). **METHODS:** We simulated the outcomes of two hypothetical treatments having three continuous covariates that are correlated with the treatments and with each other. We varied the sample size, noise levels, and tested the methods under conditions of model misspecification. To evaluate performance of the methods, we used two measures: correct identification of a statistically significant treatment effect (p < 0.05) and the root-mean-squared error for the treatment effect. **RESULTS:** For the correct-specified models, IPTW performed well relative to other methods, particularly at small sample sizes. At low noise levels and large samples sizes, all methods reliably identified a treatment effect. PSM lagged in performance for small sample sizes, and DR showed relatively weak performance under most conditions, especially under model misspecification and high noise levels. For misspecified models, the relative order of performance was similar to that of the correct-specified models. The results at high noise level were poor even for large sample sizes. **CONCLUSIONS:** MR is an unintentionally popular choice for its ease of use and the belief that covariates may adjust well for treatment effects. Our results indicate that if covariates are correlated with each other or with the treatments, one should take great care in using MR unless the sample size is large. For small sample sizes, IPTW is often the best choice even for misspecified models. PSM is a reasonable choice under low noise levels and substantial sample sizes.

RESEARCH ON METHODS – Study Design

PRM33

SQAROS – STANDARDIZED QUESTIONNAIRE TO ASSESS THE RELIABILITY OF OBSERVATIONAL STUDIES

Kaczynski L, Solnica B

Jagiellonian University Medical College, Cracow, Poland

OBJECTIVES: Health technology assessment (HTA) in accordance with evidence-based medicine (EBM) is the gold standard for determining the appropriateness of therapy funding from public funds. There is problem of getting reliable results in studies evaluating the actual effectiveness of a drug. Data on the effectiveness of health technologies are provided by the observational studies, but their credibility is low due to the lack of distinctive methodological features of randomized controlled trials. Thus, appraisal of the quality of observational studies becomes important. **METHODS:** To find the most up-to-date information on the scales for the appraisal of observational studies, systematic review in Medline using the following query: "(observational studies OR cohort studies OR case-control studies OR cross-sectional studies OR follow-up studies) AND (tool OR scale) AND (quality OR validity)" was conducted. Search time frame: till December 2013. **RESULTS:** In the review more than 100 various scales or questionnaires were found. We identified critical methodological domain, the characteristics of which determines the quality of non-randomized studies including patients selection, the impact of confounding factors, presentation of the final results or the flow rate of patients in the study. Then we developed the basic format of a questionnaire assessing the quality of observational studies, consisting of 17 questions grouped in six domains: protocol, population, intervention, end-points, results and publication. The last domain is